

OPINION

Open Access



# Lifestyle modifications: coordinating the tRNA epitranscriptome with codon bias to adapt translation during stress responses

Cheryl Chan<sup>1</sup>, Phuong Pham<sup>2</sup>, Peter C. Dedon<sup>1,2\*</sup> and Thomas J. Begley<sup>3\*</sup>

## Abstract

Cells adapt to stress by altering gene expression at multiple levels. Here, we propose a new mechanism regulating stress-dependent gene expression at the level of translation, with coordinated interplay between the tRNA epitranscriptome and biased codon usage in families of stress-response genes. In this model, auxiliary genetic information contained in synonymous codon usage enables regulation of codon-biased and functionally related transcripts by dynamic changes in the tRNA epitranscriptome. This model partly explains the association between synchronous stress-dependent epitranscriptomic marks and significant multi-codon usage skewing in families of translationally regulated transcripts. The model also predicts translational adaptation during viral infections.

## Introduction

The ‘central dogma’ defines the ‘what’ of biology—genes are transcribed into messenger RNAs that are translated into proteins. But it says nothing about the ‘when’ or ‘how much’ of gene expression. The application of systems-level ‘-omic’ technologies has led to the discovery of information-rich and combinatorial scheduling systems for gene expression involving dozens of enzyme-catalyzed chemical modifications of DNA, RNA and proteins—the epigenome and epitranscriptome. Here, we explore the evidence for a mechanism of translational control of gene expression in which the earliest and best-known RNA ‘marks’—the dozens of modified nucleosides in the transfer RNA (tRNA) epitranscriptome—interact with what amounts to families of transcripts that possess skewed usage patterns for many codons to fast-track production of survival proteins during stress.

The scheduling of gene expression at the level of transcription is well established in the literature of transcription factors, splicing, and messenger RNA (mRNA) stability, among many other mechanisms, including the complicated function of micro-RNAs (miRNAs), Piwi-

interacting RNA (piRNA), small nuclear RNA (snRNA), long non-coding RNA (lncRNA), tRNA-derived stress-induced RNA (tiRNA), and tRNA-derived RNA fragment (tRF) RNA [1]. The idea of heritable mechanisms for scheduling transcription in eukaryotic cells has more recently emerged in the form of the epigenome, with unique DNA and histone protein modification patterns in each cell type determining which genes are transcribed [2, 3]. However, there has been a long-standing dilemma posed by the observation that the correlation between changes in levels of mRNA and protein can be relatively poor, with correlation coefficients on the order of 0.4 [4, 5]. The correlation in changes in mRNA and protein levels improves somewhat when systematic delays between transcription and translation are considered [6], but other mechanisms must exist to account for changes in protein levels that do not reflect the abundance of their mRNA. Among the post-transcriptional mechanisms for scheduling gene expression, protein degradation mechanisms [7] have been more clearly delineated than translational mechanisms, with the latter largely focused on translation initiation, efficiency, and fidelity rather than scheduling [8].

\* Correspondence: [pcdedon@mit.edu](mailto:pcdedon@mit.edu); [tbegley@albany.edu](mailto:tbegley@albany.edu)

<sup>1</sup>Singapore-MIT Alliance for Research and Technology, 1 CREATE Way 138602, Singapore

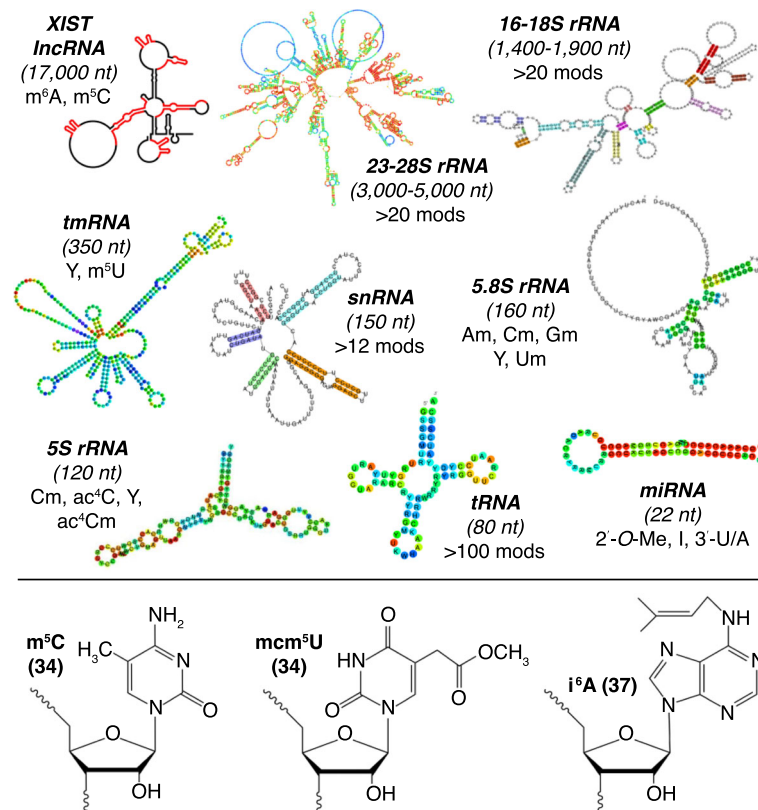
<sup>3</sup>The RNA Institute, College of Arts and Science, University at Albany, SUNY, Albany, NY 12222, USA

Full list of author information is available at the end of the article



Enter the epitranscriptome—now defined as the set of modified ribonucleotides in all forms of RNA, coding and non-coding (Fig. 1). While more than 120 different modifications have been cataloged in transfer RNA (tRNA) and ribosomal RNA (rRNA) over the past half-century, the term ‘epitranscriptome’ was first applied in 2012 to the subset of modifications in mRNA, with the observation of dynamic changes in the levels of N<sup>6</sup>-methyladenosine (m<sup>6</sup>A) in different transcripts [9]. It is now recognized that virtually every form of RNA contains modified ribonucleotides (Fig. 1), and these have been extensively reviewed elsewhere [10–16]. Significant advances have been made in defining the role of individual modifications in regulating RNA stability and translation rate, efficiency, and fidelity [10, 17–37]. However, the most mechanistically detailed models for systems-level functions of RNA modifications have arisen from studies of stress reprogramming of the tRNA epitranscriptome, and were further supported by computational studies of codon usage and the analysis of codon pausing using ribosome profiling of cells deficient in or ‘addicted’

to tRNA modification enzymes [21–23, 25, 27–29, 34, 38–40]. Based on a growing literature, we propose a model for scheduling protein synthesis of many crucial stress-response proteins involving coordinated interactions between the tRNA epitranscriptome and a select group of transcripts possessing skewed usage patterns for many codons (Fig. 2a). Supporting this ‘translational control model’, stress-induced changes in tRNA modifications that regulate the translation of codon-biased transcripts have been observed in bacteria, yeast, and mouse cells [21–23, 27–29, 34, 38, 39]. Defects in tRNA modifications and specific tRNA-modifying enzymes have also been shown to coordinately regulate the synthesis of groups of proteins from codon-biased genes in human cancer models and many other cell types [21–23, 25, 27–29, 34, 38–41]. Here, we consider the general principles of how dynamic changes in the tRNA epitranscriptome can coordinately regulate the translation of codon-biased transcripts. We note that this model is not applicable for all the expressed transcripts in a cell, but the ~5 to 10% that have statistically significant



**Fig. 1** Prototypical RNA species and their modifications. All types of RNA are modified; structures for three of more than 120 modifications are shown in the lower panel. RNA secondary structures were adapted from the following sources: XIST long non-coding RNA (lncRNA) [82]; 5S and 5.8S rRNA, tRNA, microRNA (miRNA), small nuclear RNA (snRNA), transfer-messenger RNA (tmRNA), 16-18S, 23-28S rRNA [83]. Abbreviations: ac<sup>4</sup>C N4-acetylcytidine, ac<sup>4</sup>Cm N4-acetyl-2'-O-methylcytidine, Am 2'-O-methyladenosine, Cm 2'-O-methylcytidine, Gm 2'-O-methylguanosine, I inosine, i<sup>6</sup>A N6-isopentenyladenosine, mcm<sup>5</sup>U 5-methoxycarbonylmethyluridine, m<sup>5</sup>C 5-methylcytidine, m<sup>5</sup>U 5-methyluridine, m<sup>6</sup>A N<sup>6</sup>-methyladenosine, nt nucleotide, 3'-U/A 3'-uridylation/adenylation, 2'-O-Me 2'-O-methylation, Um 2'-O-methyluridine, Y pseudouridine

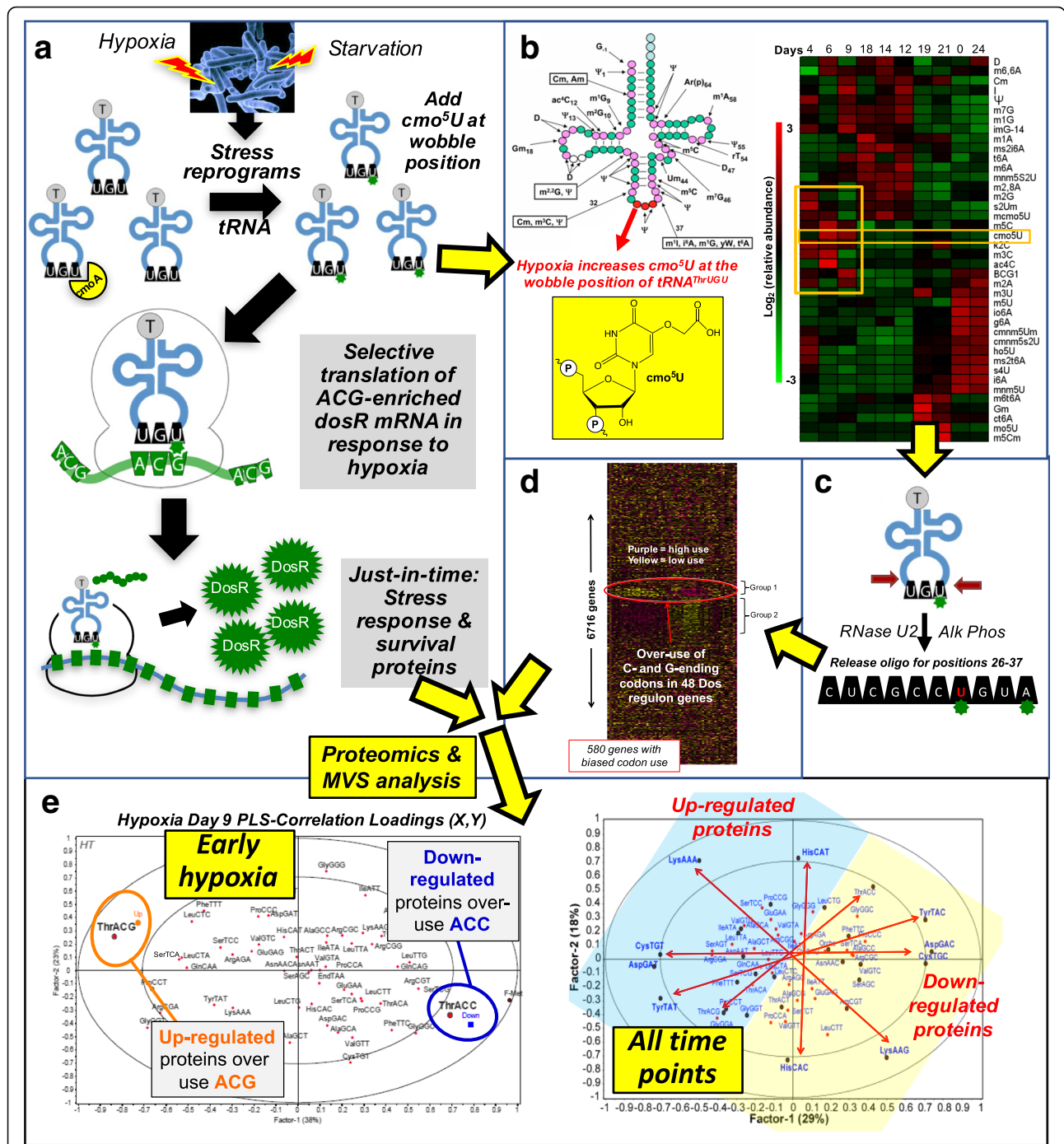


Fig. 2 (See legend on next page.)

(See figure on previous page.)

**Fig. 2** tRNA reprogramming and codon-biased translation of stress-response proteins. **a** Model illustrating the effects on *Mycobacterium bovis* BCG of hypoxia encountered during infection. Hypoxia induces expression of 48 proteins in the Dos regulon that causes the cell to become dormant. **b** The tRNA epitranscriptome consists of over 120 post-transcriptionally inserted modified ribonucleosides. During hypoxia, the relative quantities of 40 tRNA modifications (rows in heat map) change as a function of time during the response to the stress (days 0–21; columns in heat map) and again during O<sub>2</sub> resuscitation (days 22–24). In early hypoxia (day 9), the wobble position of tRNA<sup>UGU</sup>, which reads the codon ACG, switches from 5-methoxyuridine (mo<sup>5</sup>U) to 5-oxycetic acid uridine (cmo<sup>5</sup>U; structure shown). **c** RNase/LC-MS maps cmo<sup>5</sup>U to the wobble of tRNA<sup>UGU</sup>. **d** Families of response genes are organized by biased use of synonymous codons. The heat map shows over-use (purple) and under-use (yellow) of 62 codons (columns) across all genes (rows) in BCG. The gene for DosR, the master regulator of the 48-gene Dos regulon, over-uses the ACG codon and under-uses ACC, the most common Thr codon. **e** Codon analysis of proteomics data shows that > 80% of proteins upregulated in early hypoxia use ACG to code for Thr, whereas downregulated proteins are enriched in the so-called ‘optimal’ codon for Thr, ACC. Evidence that the auxiliary information in the genetic code is utilized for regulatory purposes is supported by examining codons associated with highly upregulated and downregulated proteins across all time-points of hypoxia in BCG. Pairs of synonymous codons are differentially enriched in upregulated and downregulated proteins, with the codon enrichments defining functional gene families. *Alk phos* alkaline phosphatase, *PLS* partial least squares, *Thr* threonine

deviations in codon-usage patterns for multiple codons, relative to genome averages. We primarily illustrate detailed features of our model with our work on the mycobacterial response to hypoxic stress. We conclude by describing the potential implications of the tRNA epitranscriptome and codon-usage patterns for viral infections.

### Different stresses uniquely reprogram dozens of tRNA modifications

All cells possess ~30–50 post-transcriptional modifications of ribonucleosides in RNA—known as the epitranscriptome (see Fig. 1 for sample structures and Fig. 2b for names). With 10% of the ~80 nucleotides in tRNA modified from A, C, G, and U, tRNA is the most heavily decorated form of RNA, greatly exceeding the diversity and frequency of modification levels in mRNA, rRNA, and other forms of RNA. Many of the > 120 modified ribonucleosides identified to date [42, 43] influence tRNA stability [17, 18, 20] and translational speed [30] and fidelity [19, 26, 31–33, 36]. For example, 1-methyladenosine (m<sup>1</sup>A) at position 58 of both yeast and human tRNA<sup>iMet</sup> has been reported to be crucial for its stability [17, 18], whereas ‘restrictive’ modifications at the tRNA anticodon wobble uridine position 34 in prokaryotes and eukaryotes, such as 5-methylaminomethyl-2-thio (mnm<sup>5</sup>s<sup>2</sup>) and 5-methoxycarbonylmethyl-2-thio (mcm<sup>5</sup>s<sup>2</sup>), respectively, enhance the translation of the cognate codons by enhancing ribosome A-site occupancy, stacking interactions for codon–anticodon interactions, and prevent frame shifting [19, 26, 36]. Similarly, modifications at tRNA position 37 adjacent to the anticodon loop, such as N6-threonylcarbamoyladenine (t<sup>6</sup>A) and N6-isopentenyladenine (i<sup>6</sup>A), are thought to prevent frame shifting, thereby enhancing the fidelity of translation [19, 33, 36].

An important enabler in the study of epitranscriptomics was the development of sensitive mass-spectrometry-based ‘-omic’ approaches to simultaneously identify and quantify the full set of modified ribonucleosides

[44–47], which complements published studies quantifying individual modification systems and their link to codon-biased translation [22, 23, 38, 40, 48–51]. Using a sensitive method based on liquid chromatography–tandem mass spectrometry (LC-MS/MS) [44, 52], it was found that cells respond to different stresses by uniquely changing the relative quantities of ~30–50 tRNA modifications. It was discovered that tRNA modifications behave in a coordinated manner to control gene expression at the level of translation during cell stress responses—the first facet of the translational control model (Fig. 2a). Subsequent studies profiling the tRNA modification landscape from bacteria [34] to yeast [21, 27, 39, 53] to mammalian cells [29] subjected to a variety of stressors revealed the generality of tRNA epitranscriptome reprogramming as part of the cell-stress response. For example, mycobacteria respond to hypoxic stress by gradually shifting from logarithmic growth to a non-replicative, drug-resistant dormant state over several weeks. The heat map in Fig. 2b illustrates the tRNA modification reprogramming dynamics in response to hypoxia—unique patterns of increased (red) and decreased (green) levels of ~40 tRNA modifications (rows) at each time-point (columns) during hypoxia (days 0–21) and subsequent rescue by re-aeration (days 22–24). A striking ‘oxic’ pattern of tRNA modification changes can be observed common to both pre-hypoxia (day 0) and during re-aeration (day 24) conditions, but not during hypoxia (days 4–21) [34]. These tRNA modification signatures are > 80% predictive of specific chemical stresses in yeast and hence contain significant information relevant to the underlying stress-response mechanisms [27, 39]. How this population-level information, which has been observed by others [54, 55], can be linked to mechanisms governing translation became apparent when stress-regulated modifications were mapped to specific tRNAs.



### Stress reprograms specific wobble modifications in specific tRNA isoacceptors

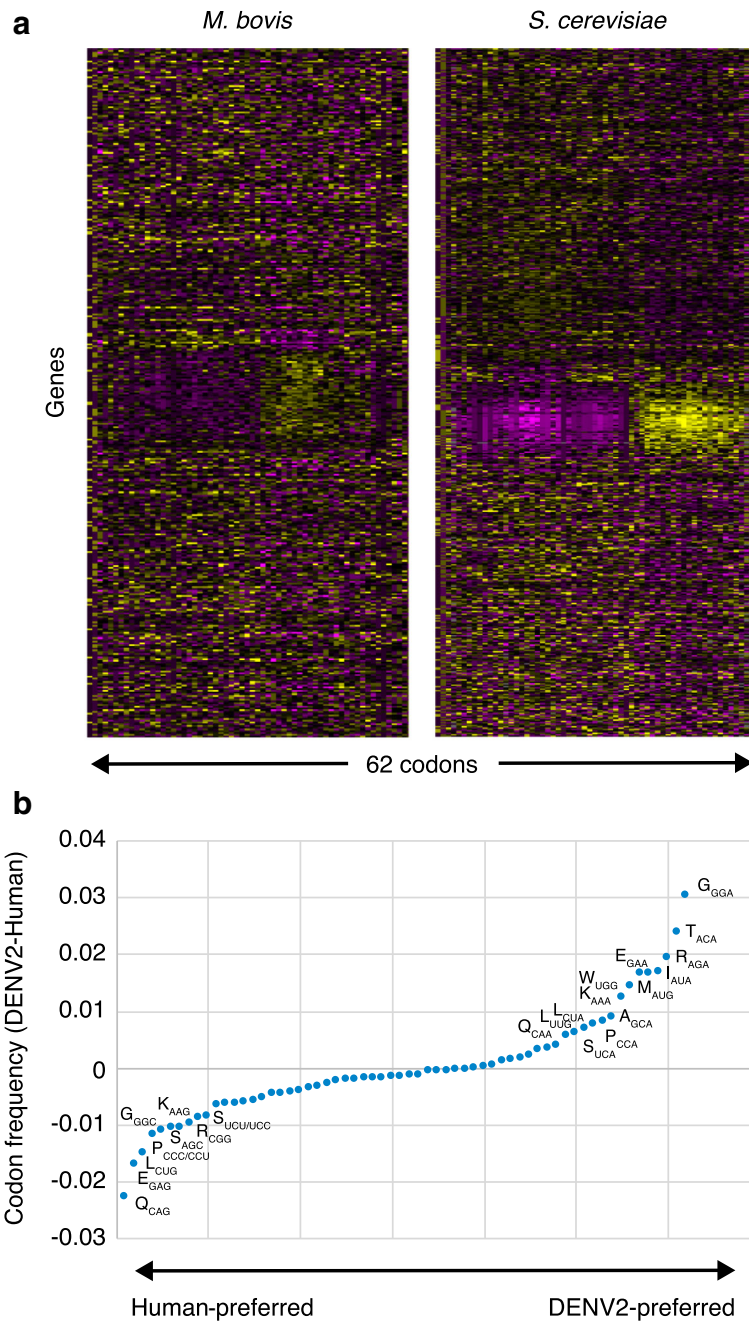
The link between stress-specific modification ‘reprogramming’ and the regulation of protein translation became clearer through mapping the stress-altered ribonucleosides to specific tRNA isoacceptors (Fig. 2c) [44]. By leveraging another mass-spectrometry (MS)-based approach similar to that applied to proteomics—MS-sequencing of specific tRNAs [44, 45]—the exact location of stress-altered ribonucleosides could be mapped in individual tRNA molecules. For example, the 5-oxycetic acid uridine (cmo<sup>5</sup>U) modification observed to increase as part of the early response to hypoxia (day 6–9) in mycobacteria (Fig. 2b, yellow box) was mapped to the wobble position of tRNA<sup>Thr</sup> that reads the ACG codon (Fig. 2c). Using a systems-level codon analytics tool [56], this observation led to a key discovery—that cognate codons of tRNAs possessing stress-altered modifications, such as cmo<sup>5</sup>U-dependent ACG, are enriched in some families of stress-response genes. Furthermore, these ACG-enriched transcripts are also over-utilizing other codons. Our findings support the idea that the auxiliary genetic information found in the form of synonymous codons is utilized during stress responses. In addition, we and others have predicted that there exists a system of differential use of subsets of the genetic code, which can also be defined as using codon bias as auxiliary genetic information [28, 34, 38, 57]. In many cases, one of the synonymous codons specific to each amino acid is differentially enriched in genes coding for proteins that are translationally regulated by epitranscriptomic marks during a specific stress (Fig. 2d, e). This same phenomenon linking stress-dependent changes in the wobble modifications of specific isoacceptors or modification enzymes to biased use of the cognate codon in stress-response genes has been observed in diverse cell types by various groups [21–23, 25, 27–29, 34, 38–40].

### Building the case for codon-biased translation—Families of stress-response genes are distinguished by biased use of synonymous codons

The idea that there is a system of biased codon usage in groups of transcripts represents the second facet of the translational control model (Fig. 2a). While obviously affected by the GC-content of the genome of an organism, the biological functions of biased codon usage have been the subject of numerous studies over the past several decades and include proposals for a relationship between codon usage in 5′-mRNA secondary structures and translation initiation, modest correlations between enrichment with ‘optimal’ codons and the abundance of cognate tRNA

isoacceptors in the pool, and differential effects of codon usage on translation elongation rate, translation fidelity and, likely, protein folding [22, 30, 58–68]. Building on the observation that genes encoding amino acid biosynthesis pathways are enriched with rare codons that are read by tRNAs that remain highly charged during amino acid starvation [64, 69], several groups have proposed the general idea of a role for biased codon usage in families of stress-response genes [21–23, 25, 27–29, 34, 38–40]. This idea was supported by gene-specific analysis of codon-usage patterns across eukaryotic and prokaryotic genomes using an algorithm to quantify usage frequencies of each of the 62 codons in each gene in an organism relative to genome-average values. The genomes of *Mycobacterium bovis* BCG and the budding yeast *Saccharomyces cerevisiae* (Fig. 3; Additional file 1) [34, 38, 56] illustrate the point that there are hundreds of genes with biased codon-usage patterns (Fig. 3, ‘over-used’ codons in yellow, ‘under-used’ codons in purple) relative to genome averages. It is important here to clarify the difference between ‘over-used’ codons and the so-called ‘optimal’ codons—codons that are the most frequently used based on a genome average, with a high abundance of the decoding tRNAs in the pool. There is a widely held view that the most abundantly expressed proteins are the result of efficient translation of genes enriched in ‘optimal’ codons that match the most abundant tRNAs in the pool [65, 66]. This has led to the current practice of codon optimization for foreign genes expressed in a different host cell, in which the codons in the foreign gene are replaced with the ‘optimal’ codons of the host cell (e.g., see [70]). However, ‘optimal’ codons only reflect very specific high-abundance proteins expressed in unstressed cells grown under optimal conditions, with the many codon-biased genes not falling into the category of being highly expressed [56, 71].

In contrast to classical codon optimality and many of the proposed functional models for genomic codon bias, a number of studies in eukaryotes and prokaryotes have now shown that groups or families of genes involved in stress responses systematically over-use and under-use specific ‘non-optimal’ synonymous codons. This is illustrated by the observation that the 48 genes in the DosR regulon in *M. bovis* BCG, which are essential for survival under hypoxic stress [34, 72], are enriched in G- and C-ending codons [34] (Figs. 2d and 3). More examples of this idea of codon-biased stress-response genes are provided by amino acid biosynthesis genes that are crucial during amino acid starvation in *Escherichia coli* [64, 69], Elongator-dependent translation of codon-biased families of cell-division genes [22, 23], and wobble uridine U34 modification-dependent regulation of glycolytic genes by codon-biased expression of the *HIF1A* gene in melanoma cells [40], as well as tRNA modification-dependent and codon-dependent regulation of the oncoprotein DEK,



**Fig. 3** (See legend on next page.)

(See figure on previous page.)

**Fig. 3** Patterns of synonymous codon usage define families of stress-response genes and might predict epitranscriptomic responses to viral infections. **a** The idea that translation regulation uses auxiliary genetic information in the form of codon bias arose by linking systems-level analyses of stress-induced proteomic upregulation and downregulation [21, 28, 34] with codon analytics [28, 38]. The heat maps shown here are examples of genome-level application of a codon-counting algorithm and visualization approaches [28, 34] to the genomes of *Mycobacterium bovis* BCG and *Saccharomyces cerevisiae*. The maps show over-use (yellow) and under-use (purple) of 62 codons (columns) across all genes (rows). For each organism, clusters represent groups of genes that have distinctly different codon-usage patterns compared with genome averages, with two opposing groups of genes identified in humans. As shown in Fig. 2d, the smaller of two clusters of codon-biased genes in *M. bovis* BCG consists of the DosR regulon of 48 genes that control the response to hypoxic stress [34]. **b** Widely differing codon-usage patterns in the human genome and a representative dengue serotype 2 genome (DENV2) could predict reprogramming of the host cell tRNA epitranscriptome both to accommodate the codon mismatch in the viral RNA genome and to respond to the stress of viral infection. Codon frequency data were generated using the web-analysis interface on the published Codon Utilization Tool (CUT) [56] and the publicly available human and dengue sequence information (Human refseq\_hg38 and dengue virus 2, complete genome NCBI Reference Sequence: KM204118.1), together with human and dengue frequency data found in Additional file 1. The analysis shows that the DENV2 genome is biased toward A-ending codons, whereas the human genome is biased toward G- and C-ending codons. This leads to the testable hypothesis that DENV2 infections will cause changes in the host cell tRNA pool, both modifications and copy numbers, to simultaneously accommodate translation of the viral mRNA and facilitate translation of codon-biased host stress-response genes

which regulates the IRES-dependent translation of the transcription factor LEF1 in breast cancer metastasis models [41]. These observations of codon bias in gene families and global regulators further support the idea that some transcripts will be more efficiently translated under specific stress conditions through stress-induced changes in the tRNA epitranscriptome in order to optimize translation and levels of the appropriate protein response systems and networks (Fig. 2). The evidence for this model is discussed below.

### Stress-reprogrammed tRNAs are required for translation of codon-biased mRNAs encoding stress-response proteins in bacteria, yeast, and mammals

That there is a mechanistic link between the stress-reprogrammed tRNA epitranscriptome, the existence of gene-specific codon-usage patterns, and selective translation of codon-biased mRNAs for stress-response genes (Fig. 2a) is borne out in a variety of studies in bacteria, yeast, and human cells [21–23, 25, 27–29, 34, 38–40]. The most striking illustrations of this mechanism arise from linked analysis of stress-induced changes in the transcriptomes, epitranscriptomes, and proteomes in yeast and bacteria [21, 28, 34]. In the case of *M. bovis* BCG subjected to hypoxic stress, multivariate statistical analysis of proteomic data revealed that pairs of synonymous codons were differentially enriched in genes for proteins that were upregulated or downregulated across the hypoxia time-course (Fig. 2e). For example, the early-response (day 9) of mycobacteria to hypoxic stress involves upregulating proteins from genes enriched with the Thr codon ACG and downregulating proteins from genes enriched with the synonymous Thr codon ACC (Fig. 2e, left panel). That the expression of codon-biased gene families is controlled mainly at the level of translation and not transcription was established by correcting the proteomic analyses for mRNA expression and protein abundance differences [21, 28, 34, 38].

We have further observed that this dichotomous pattern of differentially used codons generalizes in all sampled time-points of the mycobacteria hypoxic stress response (Fig. 2e)—under a specific stress condition, most upregulated proteins are enriched with one synonymous codon, whereas downregulated proteins are enriched with the partner synonymous codon. It is important to point out that hypoxic stress increased the translation of proteins from genes enriched with non-‘optimal’ codons ACG (Thr), CTA (Leu), GCG (Ala), and GGA (Gly), whereas their synonymous partners ACC (Thr), CTT (Leu), GCT (Ala), and GGT (Gly) were all overrepresented in downregulated proteins in hypoxia [34].

In some cases, an organism will use an optimal codon more than expected and pair it with low-usage codons in a transcript [21, 28, 38]. Alkylation stress by exposure to methyl methanesulfonate (MMS) increased Trm9-dependent wobble uridine 5-methoxycarbonylmethyl ( $mcm^5$ ) and  $mcm^5s^2$  modifications linked to tRNA<sup>Arg</sup>(UCU) and tRNA<sup>Glu</sup>(UUC), with increased translation of mRNAs enriched in the cognate AGA and GAA codons, respectively [27, 38, 53]. Interestingly, however, the Trm9-dependent codons AGA (Arg) and GAA (Glu) are enriched in DNA damage and cell-cycle control genes crucial to surviving alkylation stresses [28, 38]. In addition, the non-optimal codons GAC (Asp), ATC (Iso), TAC (Tyr), AAG (Lys), and TTC (Phe) are found paired with AGA and GAA on the transcripts whose translation is linked to 5-methoxycarbonylmethyluridine ( $mcm^5U$ ) and 5-methoxycarbonylmethyl-2-thiouridine ( $mcm^5s^2U$ ) [38]. The identified codon-biased genes are among those clustering in the heat map of codon usage in the *S. cerevisiae* genome shown in Fig. 3. Coupled with the biased use of synonymous codons in families of stress-response genes, this more than pair-wise use of one synonymous codon from a set to control protein upregulation and downregulation represents a systematic repurposing of the auxiliary information tied

to the genetic code for adaptation and survival in a changing environment.

This association between the stress-altered epitranscriptome and translation of codon-biased stress-response genes has also been observed in two forms of yeast, the nematode *Caenorhabditis elegans*, and human [21–23, 25, 27–29, 34, 38–40]. For example, *S. cerevisiae* responded to H<sub>2</sub>O<sub>2</sub>-induced oxidative stress by increasing Trm4 methyltransferase-dependent wobble 5-methylcytidine (m<sup>5</sup>C)-modification in tRNA<sup>Leu</sup>(CAA), which resulted in the selective translation of mRNA from genes enriched in the cognate TTG codon [21]. Similarly, mitosis and cytokinesis in the fission yeast *Schizosaccharomyces pombe* is controlled by Elongator-mediated synthesis of mcm<sup>5</sup>s<sup>2</sup> in tRNAs recognizing AA-ending codons, with these codons enriched in three different groups of proteins, including proteins involved in cell division [22, 23]. More recently, Close and colleagues demonstrated that the carcinogenic and drug-resistant phenotype of human melanoma cells required Elongator-dependent translation of codon-biased mRNAs, including the stress-response mediator *HIF1A* [40]. These results are consistent with a general mechanism for enhancing or scheduling translation of proteins needed by the cell to mount an appropriate response to specific stress conditions.

Further support that stress-regulated tRNA modifications are directly linked to expression of codon-biased survival proteins has emerged in the previously described studies in yeast, *C. elegans*, and human [21–23, 25, 27–29, 38, 40]. In all cases, cells lacking the tRNA-modifying enzymes did not show the specific, modification-dependent codon-biased translation and had distinct proliferative defects [21–23, 25, 27–29, 38, 40]. In yeast, this rendered the cells hypersensitive to killing by the stresses [21, 28, 39], whereas in human melanoma cells, this reduced the tumorigenicity and drug resistance of the cancer cells [40]. For example, loss of Trm4 in *S. cerevisiae* prevented m<sup>5</sup>C formation at the wobble position of tRNA<sup>Leu</sup>(CAA), abolished selective translation of mRNA from genes enriched in the TTG codon, and resulted in hypersensitivity of cells to H<sub>2</sub>O<sub>2</sub> [21]. Similarly, deletion of Trm9 prevented formation of its product mcm<sup>5</sup>U in tRNA<sup>Arg</sup>(UCU), with a concomitant enhanced sensitivity to MMS [28, 38, 67].

In mammals, a good example of a specific tRNA epitranscriptomic mark regulating the translation of codon-biased mRNAs has been identified in the synthesis of the stress-important selenoproteins [29], which include the H<sub>2</sub>O<sub>2</sub>-detoxifying glutathione peroxidase (Gpx) and thioredoxin reductase (TrxR) enzymes. The approximately 25 known selenoproteins contain the 21st amino acid selenocysteine (Sec). Notably, Sec does not contain a dedicated codon, but instead uses a recoded stop-codon (UGA) along with specific wobble

U modifications in tRNA<sup>SER(SEC)</sup> and many other factors to promote the incorporation of Sec into the growing peptide chain [73–75]. Stress-responsive selenoproteins are an excellent example of our translational control model, albeit very specific, as corresponding transcripts have at least twice as many stop codons as expected (i.e., biased codon usage) and require increased 5-methoxycarbonylmethyl-2'-O-methyluridine (mcm<sup>5</sup>Um) modifications in tRNA to increase the levels of some Gpx and TrxR enzymes in response to H<sub>2</sub>O<sub>2</sub> [29, 56]. The observations of Close and coworkers in human melanoma cells and breast cancer cells provide two examples of mechanisms where tRNA modification-enzyme-dependent codon-biased translation of a master regulator controls network-based responses [40, 41]. Although the observations of codon-dependent translational control of HIF1 $\alpha$  and DEK are not examples of broad translational control itself, they do nevertheless control broad transcriptional networks. Combined, these findings further support the notion of coordinated interactions between the tRNA epitranscriptome and biased codon usage to enhance translation of survival proteins.

### Implications of a system linking the epitranscriptome to a code of codons—Viral infections

One example of how the proposed epitranscriptome-based gene regulation model can be ‘translated’ to other arenas involves viral infections. Viruses depend on the translational machinery of the host cell for survival and replication, which poses a significant problem in light of the highly different codon-usage patterns in the viral and host-cell genomes. Consider, for example, RNA-based Flaviviruses that rely on the translational machinery of the host for immediate synthesis of viral proteins essential for replication and survival after infection. A comparison of the codon usage frequencies in the human genome and the ~11 kb genome of a representative Flavivirus—dengue virus serotype 2 (DENV2)—showed that the viral RNA genome is heavily biased toward A-ending codons, whereas the human genome is biased toward C- and G-ending codons (Fig. 3b), in spite of the similar GC content of the two genomes (human 42% GC versus DENV2 46% GC). This raises a potential problem for the virus, in part when it needs to translate what are rare codons for the host and for which the cognate tRNAs are present in the pool at low levels or with inappropriate modifications. This mismatch between codon usage and the tRNA pool could lead to stalled translation, mistranslation, and proteotoxic stress, which would be detrimental to the virus. In addition to potentially optimizing the levels of specific host tRNA isoacceptors as a strategy by the virus to ensure efficient translation of its unique genome, our translational control



model predicts the need to reprogram the host tRNA epitranscriptome, particularly at the wobble uridine bases to enhance decoding of what would normally be considered rare A-ending codons for the human host.

An intriguing feature of the dengue genome raises questions about the ability of the virus to co-opt the host translational machinery to facilitate translation of viral proteins in the face of codon mismatch. Among the ten proteins coded by the dengue genome, the NS5 protein is an RNA methyltransferase that catalyzes formation of 7-methylguanosine ( $m^7G$ ) and 2'-O-methyladenosine (Am) in the GpppA cap added to the viral RNA, and additionally forms Am throughout the viral genome [76]. In vitro reactions revealed that NS5 can methylate human rRNA [76], raising the possibility of a viral methyltransferase that could directly coordinate reprogramming of the host translational epitranscriptome to enhance translation of the viral genome. Recent studies using a similar LC-MS/MS approach to ours but applied instead to the analysis of non-size-selected, hydrolyzed total cellular RNA and viral RNA isolated from RNA-virus-infected cell cultures showed changes in RNA post-transcriptional modifications (PTMs) occurring following viral stress on host cells [54, 77]. Future studies are needed to carefully deconvolute this initial landscape of PTMs in terms of their site, function, and interplay in the cell-stress response. Additionally, the promiscuity of the dengue NS5 methyltransferase for both rRNA and tRNA substrates is reminiscent of known dual-specificity RNA-modifying enzymes such as pseudouridine synthase RluA and RNA methyltransferase RlmN in *E. coli* that modify both tRNA and rRNA bases [78, 79]. Could this be a common mechanism to synchronize changes across epitranscriptomes? Recent work demonstrating the synchrony between cytosolic and mitochondrial translation [80] raises yet another interesting possibility of crosstalk between both epitranscriptomes that can build on the translation control model we present here. With a direct-acting methyltransferase and known modulatory effects on the cellular metabolism of the host [81], the Flavivirus–human host-cell infection model presents an attractive system to probe the ribonucleome and epitranscriptome reprogramming of the host cell in response to viral stress and to gain insights into potentially important mechanisms.

### Concluding remarks

Our proposed mechanism for translational adaptation involving coordinated interplay between the tRNA epitranscriptome and biased codon usage represents a complicated interaction among diverse systems and is well supported by observations in prokaryotes and eukaryotes. In addition to generating numerous testable hypotheses concerning controlling gene expression at the level of translation—such as using codon bias as a

predictor of epitranscriptome dynamics—this system has important implications for synthetic biology in the form of genetic tools to tune the pool of tRNA molecules and the dozens of programmable tRNA modifications, for predicting translational adaptation during viral infections, and for expression of foreign genes in cells.

### Additional file

**Additional file 1: Table S1.** Tabulated codon usage frequencies for dengue virus and human genomes. (XLSX 15 kb)

### Funding

The authors acknowledge funding by NIH grants CA026731 (PCD), ES024615 (TJB, PCD), ES026856 (TJB, PCD), and ES002109 (PCD) and the National Research Foundation of Singapore through the Singapore-MIT Alliance for Research and Technology (PCD).

### Authors' contributions

All authors read and approved the final manuscript.

### Competing interests

The authors declare that they have no competing interests.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Author details

<sup>1</sup>Singapore-MIT Alliance for Research and Technology, 1 CREATE Way 138602, Singapore. <sup>2</sup>Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. <sup>3</sup>The RNA Institute, College of Arts and Science, University at Albany, SUNY, Albany, NY 12222, USA.

Published online: 27 December 2018

### References

- Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet.* 2008;9:102–14.
- Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet.* 2008;9:465–76.
- Soshnev AA, Josefowicz SZ, Allis CD. Greater than the sum of parts: complexity of the dynamic epigenome. *Mol Cell.* 2016;62:681–94.
- Schwanhauser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global quantification of mammalian gene expression control. *Nature.* 2011;473:337–42.
- Schwanhauser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Corrigendum: global quantification of mammalian gene expression control. *Nature.* 2013;495:126–7.
- Gedeon T, Bokes P. Delayed protein synthesis reduces the correlation between mRNA and protein fluctuations. *Biophys J.* 2012;103:377–85.
- Ciechanover A. Intracellular protein degradation: from a vague idea through the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting. *Bioorg Med Chem.* 2013;21:3400–10.
- Gingold H, Pilpel Y. Determinants of translation efficiency and accuracy. *Mol Syst Biol.* 2011;7:481.
- Meyer KD, Jaffrey SR. The dynamic epitranscriptome: N6-methyladenosine and gene expression control. *Nat Rev Mol Cell Biol.* 2014;15:313–26.
- Nachtergaele S, He C. The emerging biology of RNA post-transcriptional modifications. *RNA Biol.* 2017;14:156–63.
- Koh CS, Sarin LP. Transfer RNA modification and infection - implications for pathogenicity and host responses. *Biochim Biophys Acta.* 2018;1861:419–32.
- Duechler M, Leszczynska G, Sochacka E, Nawrot B. Nucleoside modifications in the regulation of gene expression: focus on tRNA. *Cell Mol Life Sci.* 2016; 73:3075–95.

13. Hsu PJ, Shi H, He C. Epitranscriptomic influences on development and disease. *Genome Biol.* 2017;18:197.
14. du Toit ARNA. Expanding the mRNA epitranscriptome. *Nat Rev Mol Cell Biol.* 2016;17:201.
15. Jacob R, Zander S, Gutschner T. The dark side of the epitranscriptome: chemical modifications in long non-coding RNAs. *Int J Mol Sci.* 2017;18:E2387.
16. Tuorto F, Lyko F. Genome recoding by tRNA modifications. *Open Biol.* 2016; 6:160287.
17. Anderson J, Phan L, Cuesta R, Carlson BA, Pak M, Asano K, et al. The essential Gcd10p-Gcd14p nuclear complex is required for 1-methyladenosine modification and maturation of initiator methionyl-tRNA. *Genes Dev.* 1998;12:3650–62.
18. Ozanick S, Krecic A, Andersland J, Anderson JT. The bipartite structure of the tRNA m1A58 methyltransferase from *S. cerevisiae* is conserved in humans. *RNA.* 2005;11:1281–90.
19. Durant PC, Bajji AC, Sundaram M, Kumar RK, Davis DR. Structural effects of hypermodified nucleosides in the *Escherichia coli* and human tRNA<sup>Lys</sup> anticodon loop: the effect of nucleosides s2U, mcm5U, mcm5s2U, mnm5s2U, t6A, and ms2t6A. *Biochemistry.* 2005;44:8078–89.
20. Alexandrov A, Chernyakov I, Gu W, Hiley SL, Hughes TR, Grayhack EJ, Phizicky EM. Rapid tRNA decay can result from lack of nonessential modifications. *Mol Cell.* 2006;21:87–96.
21. Chan CT, Pang YL, Deng W, Babu IR, Dyavaiah M, Begley TJ, et al. Reprogramming of tRNA modifications controls the oxidative stress response by codon-biased translation of proteins. *Nat Commun.* 2012;3:937.
22. Bauer F, Hermand D. A coordinated codon-dependent regulation of translation by Elongator. *Cell Cycle.* 2012;11:4524–9.
23. Bauer F, Matsuyama A, Candriacci J, Dieu M, Scheliga J, Wolf DA, et al. Translational control of cell division by Elongator. *Cell Rep.* 2012;1:424–33.
24. Novoa EM, Pavon-Eternod M, Pan T, Ribas de Pouplana L. A role for tRNA modifications in genome structure and codon usage. *Cell.* 2012;149:202–13.
25. Fernandez-Vazquez J, Vargas-Perez I, Sanso M, Buhne K, Carmona M, Paulo E, et al. Modification of tRNA(Lys) UUU by elongator is essential for efficient translation of stress mRNAs. *PLoS Genet.* 2013;9:e1003647.
26. Rezgui VA, Tyagi K, Ranjan N, Konevega AL, Mittelstaet J, Rodnina MV, et al. tRNA tKUUU, tQUUG, and tEUUC wobble position modifications fine-tune protein translation by promoting ribosome A-site binding. *Proc Natl Acad Sci U S A.* 2013;110:12289–94.
27. Chan CT, Deng W, Li F, DeMott MS, Babu IR, Begley TJ, et al. Highly predictive reprogramming of tRNA modifications is linked to selective expression of codon-biased genes. *Chem Res Toxicol.* 2015;28:978–88.
28. Deng W, Babu IR, Su D, Yin S, Begley TJ, Dedon PC. Trm9-catalyzed tRNA modifications regulate global protein expression by codon-biased translation. *PLoS Genet.* 2015;11:e1005706.
29. Endres L, Begley U, Clark R, Gu C, Dziergowska A, Malkiewicz A, et al. Alkbh8 regulates selenocysteine-protein expression to protect against reactive oxygen species damage. *PLoS One.* 2015;10:e0131335.
30. Nedialkova DD, Leidel SA. Optimization of codon translation rates via tRNA modifications maintains proteome integrity. *Cell.* 2015;161:1606–18.
31. Zaborse JM, Bauer DuMont VL, Wallace EW, Pan T, Aquadro CF, et al. Correction: a nutrient-driven tRNA modification alters translational fidelity and genome-wide protein coding across an animal genus. *PLoS Biol.* 2015;13:e1002150.
32. Zaborse JM, DuMont VL, Wallace EW, Pan T, Aquadro CF, Drummond DA. A nutrient-driven tRNA modification alters translational fidelity and genome-wide protein coding across an animal genus. *PLoS Biol.* 2014;12:e1002015.
33. Lamichhane TN, Blewett NH, Crawford AK, Cherkasova VA, Iben JR, Begley TJ, et al. Erratum for Lamichhane et al., lack of tRNA modification isopentenyl-A37 alters mRNA decoding and causes metabolic deficiencies in fission yeast. *Mol Cell Biol.* 2015;35:1477.
34. Chionh YH, McBee M, Babu IR, Hia F, Lin W, Zhao W, et al. tRNA-mediated codon-biased translation in mycobacterial hypoxic persistence. *Nat Commun.* 2016;7:13302.
35. Grosjean H, Westhof E. An integrated, structure- and energy-based view of the genetic code. *Nucleic Acids Res.* 2016;44:8020–40.
36. Klassen R, Bruch A, Schaffrath R. Independent suppression of ribosomal +1 frameshifts by different tRNA anticodon loop modifications. *RNA Biol.* 2017; 14:1252–9.
37. Patil DP, Pickering BF, Jaffrey SR. Reading m(6)a in the transcriptome: m(6)A-binding proteins. *Trends Cell Biol.* 2018;28:113–27.
38. Begley U, Dyavaiah M, Patil A, Rooney JP, DiRenzo D, Young CM, et al. Trm9-catalyzed tRNA modifications link translation to the DNA damage response. *Mol Cell.* 2007;28:860–70.
39. Chan CT, Dyavaiah M, DeMott MS, Taghizadeh K, Dedon PC, Begley TJ. A quantitative systems approach reveals dynamic control of tRNA modifications during cellular stress. *PLoS Genet.* 2010;6:e1001247.
40. Rapino F, Delaunay S, Rambow F, Zhou Z, Tharun L, De Tullio P, et al. Codon-specific translation reprogramming promotes resistance to targeted therapy. *Nature.* 2018;558:605–9.
41. Delaunay S, Rapino F, Tharun L, Zhou Z, Heukamp L, Termathe M, et al. Elp3 links tRNA modification to IRES-dependent translation of LEF1 to sustain metastasis in breast cancer. *J Exp Med.* 2016;213:2503–23.
42. Machnicka MA, Milanowska K, Osman Oglou O, Purta E, Kurkowska M, et al. MODOMICS: a database of RNA modification pathways—2013 update. *Nucleic Acids Res.* 2013;41:D262–7.
43. Cantara WA, Crain PF, Rozenski J, McCloskey JA, Harris KA, Zhang X, et al. The RNA modification database, RNAMDB: 2011 update. *Nucleic Acids Res.* 2011;39:D195–201.
44. Cai WM, Chionh YH, Hia F, Gu C, Kellner S, McBee ME, et al. A platform for discovery and quantification of modified ribonucleosides in RNA: application to stress-induced reprogramming of tRNA modifications. *Methods Enzymol.* 2015;560:29–71.
45. Ross R, Cao X, Yu N, Limbach PA. Sequence mapping of transfer RNA chemical modifications by liquid chromatography tandem mass spectrometry. *Methods.* 2016;107:73–8.
46. Thuring K, Schmid K, Keller P, Helm M. Analysis of RNA modifications by liquid chromatography-tandem mass spectrometry. *Methods.* 2016;107:48–56.
47. Suzuki T, Ikeuchi Y, Noma A, Sakaguchi Y. Mass spectrometric identification and characterization of RNA-modifying enzymes. *Methods Enzymol.* 2007; 425:211–29.
48. Damon JR, Pincus D, Ploegh HL. tRNA thiolation links translation to stress responses in *Saccharomyces cerevisiae*. *Mol Biol Cell.* 2015;26:270–82.
49. Laxman S, Sutter BM, Wu X, Kumar S, Guo X, Trudgian DC, et al. Sulfur amino acids regulate translational capacity and metabolic homeostasis through modulation of tRNA thiolation. *Cell.* 2013;154:416–29.
50. Goodarzi H, Nguyen HCB, Zhang S, Dill BD, Molina H, Tavazoie SF. Modulated expression of specific tRNAs drives gene expression and cancer progression. *Cell.* 2016;165:1416–27.
51. Goffena J, Lefcort F, Zhang Y, Lehmann E, Chaverra M, Felig J, et al. Elongator and codon bias regulate protein levels in mammalian peripheral neurons. *Nat Commun.* 2018;9:889.
52. Su D, Chan CT, Gu C, Lim KS, Chionh YH, McBee ME, et al. Quantitative analysis of ribonucleoside modifications in tRNA by HPLC-coupled mass spectrometry. *Nat Protoc.* 2014;9:828–41.
53. Patil A, Dyavaiah M, Joseph F, Rooney JP, Chan CT, Dedon PC, et al. Increased tRNA modification and gene-specific codon usage regulate cell cycle progression during the DNA damage response. *Cell Cycle.* 2012;11:3656–65.
54. Rose RE, Pazos MA 2nd, Curcio MJ, Fabris D. Global epitranscriptomics profiling of RNA post-transcriptional modifications as an effective tool for investigating the epitranscriptomics of stress response. *Mol Cell Proteomics.* 2016;15:932–44.
55. Basanta-Sanchez M, Temple S, Ansari SA, D'Amico A, Agris PF. Attomole quantification and global profile of RNA modifications: epitranscriptome of human neural stem cells. *Nucleic Acids Res.* 2016;44:e26.
56. Doyle F, Leonardi A, Endres L, Tenenbaum SA, Dedon PC, Begley TJ. Gene- and genome-based analysis of significant codon patterns in yeast, rat and mice genomes with the CUT codon UTILization tool. *Methods.* 2016;107:98–109.
57. Maraia RJ, Iben JR. Different types of secondary information in the genetic code. *RNA.* 2014;20:977–84.
58. Bahir I, Fromer M, Prat Y, Linal M. Viral adaptation to host: a proteome-based analysis of codon usage and amino acid preferences. *Mol Syst Biol.* 2009;5:311.
59. Chaney JL, Clark PL. Roles for synonymous codon usage in protein biogenesis. *Annu Rev Biophys.* 2015;44:143–66.
60. Chiari Y, Dion K, Colborn J, Parmakelis A, Powell JR. On the possible role of tRNA base modifications in the evolution of codon usage: queuosine and drosophila. *J Mol Evol.* 2010;70:339–45.
61. Kliman RM, Irving N, Santiago M. Selection conflicts, gene expression, and codon usage trends in yeast. *J Mol Evol.* 2003;57:98–109.

62. Nakamura Y, Gojobori T, Ikemura T. Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res.* 2000;28:292.
63. Qian W, Yang JR, Pearson NM, Maclean C, Zhang J. Balanced codon usage optimizes eukaryotic translational efficiency. *PLoS Genet.* 2012;8:e1002603.
64. Quax TE, Claassens NJ, Soll D, van der Oost J. Codon bias as a means to fine-tune gene expression. *Mol Cell.* 2015;59:149–61.
65. Sharp PM, Bailes E, Grocock RJ, Peden JF, Sockett RE. Variation in the strength of selected codon usage bias among bacteria. *Nucleic Acids Res.* 2005;33:1141–53.
66. Sharp PM, Li WH. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* 1987;15:1281–95.
67. Patil A, Chan CT, Dyavaiah M, Rooney JP, Dedon PC, Begley TJ. Translational infidelity-induced protein stress results from a deficiency in Trm9-catalyzed tRNA modifications. *RNA Biol.* 2012;9:990–1001.
68. Tyagi K, Pedrioli PG. Protein degradation and dynamic tRNA thiolation fine-tune translation at elevated temperatures. *Nucleic Acids Res.* 2015;43:4701–12.
69. Dittmar KA, Sorensen MA, Elf J, Ehrenberg M, Pan T. Selective charging of tRNA isoacceptors induced by amino-acid starvation. *EMBO Rep.* 2005;6:151–7.
70. Chin JX, Chung BK, Lee DY. Codon optimization OnLine (COOL): a web-based multi-objective optimization platform for synthetic gene design. *Bioinformatics.* 2014;30:2210–2.
71. Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, et al. Global analysis of protein expression in yeast. *Nature.* 2003;425:737–41.
72. Boon C, Dick T. How *Mycobacterium tuberculosis* goes to sleep: the dormancy survival regulator DosR a decade later. *Future Microbiol.* 2012;7:513–8.
73. Gladyshev VN, Hatfield DL. Selenocysteine-containing proteins in mammals. *J Biomed Sci.* 1999;6:151–60.
74. Korotkov KV, Novoselov SV, Hatfield DL, Gladyshev VN. Mammalian selenoprotein in which selenocysteine (sec) incorporation is supported by a new form of sec insertion sequence element. *Mol Cell Biol.* 2002;22:1402–11.
75. Kryukov GV, Castellano S, Novoselov SV, Lobanov AV, Zehrab O, Guigo R, et al. Characterization of mammalian selenoproteomes. *Science.* 2003;300:1439–43.
76. Dong H, Chang DC, Hua MH, Lim SP, Chionh YH, Hia F, et al. 2'-O methylation of internal adenosine by flavivirus NS5 methyltransferase. *PLoS Pathog.* 2012;8:e1002642.
77. McIntyre W, Netzband R, Bonenfant G, Biegel JM, Miller C, Fuchs G, et al. Positive-sense RNA viruses reveal the complexity and dynamics of the cellular and viral epitranscriptomes during infection. *Nucleic Acids Res.* 2018; 46:5776–91.
78. Wrzesinski J, Nurse K, Bakin A, Lane BG, Ofengand J. A dual-specificity pseudouridine synthase: an *Escherichia coli* synthase purified and cloned on the basis of its specificity for psi 746 in 23S RNA is also specific for psi 32 in tRNA(phe). *RNA.* 1995;1:437–48.
79. Benitez-Paez A, Villarroya M, Armengod ME. The *Escherichia coli* RlmN methyltransferase is a dual-specificity enzyme that modifies both rRNA and tRNA and controls translational accuracy. *RNA.* 2012;18:1783–95.
80. Couvillion MT, Soto IC, Shipkovenska G, Churchman LS. Synchronized mitochondrial and cytosolic translation programs. *Nature.* 2016;533:499–503.
81. Jordan TX, Randall G. Flavivirus modulation of cellular metabolism. *Curr Opin Virol.* 2016;19:7–10.
82. Fang R, Moss WN, Rutenberg-Schoenberg M, Simon MD. Probing Xist RNA structure in cells using targeted structure-seq. *PLoS Genet.* 2015;11: e1005668.
83. Hubley R, Finn RD, Clements J, Eddy SR, Jones TA, Bao W, et al. The Dfam database of repetitive DNA families. *Nucleic Acids Res.* 2016;44:D81–9.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

